JUAN RAMIREZ

ramirezjuan9806@gmail.com +57 3209161892

Medellín, Colombia https://www.linkedin.com/in/juan-esteban-ramirez-reyes/ https://github.com/juan8991

SUMMARY

Mid-senior Backend Engineer with over 3 years of experience in Java and Python, using frameworks like Spring Boot, Spring WebFlux, and FastAPI, specialized in microservice architectures deployed on-premise and Cloud with providers like AWS and GCP, with a growing specialization in AI agent systems and LLM integrations (Gemini, OpenAI, Claude), delivering significant performance improvements including reduced API response times from 1.8s to 300ms, increased order processing capacity from 300 to 5,000+ per day, an 85% reduction in invoice processing time (6h to 50min) for 25K+ monthly records using a Spring Batch ETL pipeline, and multi-agent systems capable of qualifying 500+ leads/day.

SKILLS

Languages & Frameworks: Java, Python, TypeScript, Spring Boot, Spring Batch, FastAPI, Next.js, Angular.
Databases: PostgreSQL, MySQL, MariaDB, MongoDB, Redis, SQLite, RDS, Supabase.
Integration Platforms : N8N (self-hosted), GoHighLevel, Evolution API (self-hosted), Dapta, ElevenLabs.
CI/CD & Testing : GitHub Actions, GitLab CI/CD, JUnit, Mockito, Selenium, SonarQube.
Cloud & Infrastructure : AWS (EC2, ECS, EKS, Lambda, S3), GCP, Docker, Linux, Nginx, Apache.
AI Tools : Claude Code, Codex, Open Code, Antigravity, Herdrz, Orca, Google AI Studio, Google Flow.

WORK EXPERIENCE

Mid-senior Backend Engineer, Auratech, Medellin, Colombia. *Aug 2025 - Jul 2026*

- Designed and developed a platform to automate the core business operations, enabling a 24/7 uninterrupted system. This reduced operational workload from 10 to 3 people and increased revenue by over $200M COP, using Java, Python, N8N and Angular 20.
- Reduced invoice processing time by 85% (from 6 hours to 50 minutes) for 25K+ monthly records by implementing a Spring Batch ETL pipeline with partitioning, chunk-oriented processing, and parallel step execution, handling complex validations and multi-source data aggregation
- Built an end-to-end AI-powered sales and customer engagement system for a real estate company by integrating GoHighLevel with Evolution API, N8N (as middleware), WhatsApp, Instagram, Facebook, and TikTok, enabling 24/7 lead qualification, material delivery, and appointment scheduling with less than 30 second response times, capable of handling 500+ leads/day and booking 10+ qualified appointments per week.
- Architected a multi-agent AI system using OpenClaw, Gemini, and Claude, deploying 4 specialized personal assistants for area leaders (accounting, pharmaceutical, data intelligence, and general management) plus a sales agent with voice note generation (ElevenLabs), image and document delivery capabilities; built a real-time Mission Control dashboard in Next.js and Supabase to visualize agent interactions and customer flows.
- Developed a legal document automation platform in Python and Next.js integrated with Dapta (voice agent) and WhatsApp, automating the generation of PQRs and Tutelas (constitutional legal filings in Colombia), achieving a 53% increase in daily document processing capacity.

Mid-senior Backend Engineer, Aunna It, Medellin, Colombia. *Jul 2024 - Jul 2025*

- Directed the development of an automated system for the search, approval, and delivery of medications, scaling operational capacity from hundreds to thousands of medications processed per day using PHP, Python, Java, Vertex AI, and Vue.js.
- Led the migration of legacy systems from Node.js and AngularJS to Angular 16 and PHP with Laravel, improving performance, security, and scalability of the application.
- Led the restructuring and stabilization of a critical production project, eliminating client losses and scaling the platform to support thousands of daily users through core business optimization and database query tuning with PHP and Vue.js.
- Implemented automated CI/CD pipelines for legacy projects under maintenance using GitLab and Runners on Linux, which eliminated 100% of human error during deployments and drastically reduced delivery time.

Backend Engineer, Aunna It, Medellin, Colombia. *June 2023 - June 2024*

- Achieved sub-200ms response times (from 2s average) and improved system maintainability by migrating to a scalable Java microservices architecture.
- Increased query efficiency by 50%, reducing average execution time from 1.2 seconds to 600ms through MariaDB relational database optimization with normalization and strategic indexing.
- Reduced monthly downtime from 5 hours to 1 hour (80% decrease) by leading critical issue resolution through infrastructure and code improvements, ensuring higher system reliability using Java and PHP.
- Collaborated on modernizing legacy systems, decreasing monthly support hours by 48% (from ~25h to ~13h).

Backend Engineer, Alternova, Medellin. *Nov 2022 - May 2023*

- Accelerated partner onboarding time-to-market by 75% (from 4 weeks to 1 week) through the development of a reusable catalog of RESTful APIs and SOAP services, improving interoperability across the application ecosystem.
- Reduced response time of critical queries by 40% (from 2.5s to 1.5s) in high-volume MySQL databases by implementing advanced indexing and normalization techniques, minimizing latency and optimizing resource usage in intensive operations.
- Improved code quality and maintainability by implementing a unit testing strategy across 10+ microservices, achieving 90% coverage and reducing production bugs by 18%.

EDUCATION

Universidad del Rosario - Escuela Colombiana de Ingenieria Julio Garavito, Bogota, Colombia.
Bachelor of Biomedical Engineering. *Jan 2016 - Apr 2022*

CERTIFICATIONS

Google - Associate Cloud Engineer

OTHER SKILLS

Awards: Winner Hackathon SIC 2023.